

### A. Some Lemmas

In order to prove Theorem 3, we need to introduce the following Lemmas 1 ~ 6.

**Lemma 1.** (Lemma 1 in [25]): For nonnegative numbers  $a, b, c, d, x, y$ ,

$$(ax + by)^2 + (cx + dy)^2 \leq \left( \sqrt{a^2 + c^2}x + (b + d)y \right)^2. \quad (16)$$

**Lemma 2.** Consider the general CS model  $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{E}$ ,  $\text{supp}(\mathbf{X}) = T$  and  $|T| = s$ . Suppose  $S, T_0 \subseteq \{1, 2, \dots, n\}$ ,  $|S| = t$ .  $\mathbf{W}_{T_0}$  is constructed with diagonal entries indexed by  $T_0$  being  $\omega \geq 0$ . Let  $\tilde{\mathbf{X}}$  be the solution of the least squares problem  $\arg \min_{\mathbf{Z}} \{\|\mathbf{Y} - \mathbf{A}\mathbf{Z}\|_2, \text{supp}(\mathbf{Z}) \subseteq S\}$ . Let  $s_1 := \min_{i,k} |\Lambda_{S_i}^k \cap \text{supp}(\mathbf{X}_{S_i})|$ ,  $s_2 := \min_{i,k} |\Lambda_{S_i}^k \cap \text{supp}(\mathbf{X}_{S_i}) \cap \tilde{S}_{S_i}^k|$ ,  $s_3 := \min_{i,k} |\Lambda_{S_i}^k \cap \text{supp}(\mathbf{X}_{S_i}) \cap \tilde{S}_{S_i}^k \setminus S_{S_i}^k|$  and  $\eta = \min_{i \in \{1,2,\dots,r\}} \min_{j \in \{1,2,\dots,n\}} \|(\mathbf{X}_{S_i})_{j,:}\|_2 / \|\mathbf{X}_{S_i}\|_F$ . Let  $\nu_1 = (1 + \omega)/(1 + \eta\omega\sqrt{s_2})$  and  $\nu_2 = (1 + \omega)/(1 + \eta\omega\sqrt{s_3})$ , if  $\delta_{3s} < 1$ , then

$$\|\mathbf{W}_{T_0}(\mathbf{X} - \tilde{\mathbf{X}})_S\|_F \leq \omega\delta_{s+t} \|\mathbf{X} - \tilde{\mathbf{X}}\|_F + \omega\sqrt{1 + \delta_t} \|\mathbf{E}\|_F \quad (17)$$

and

$$\|\mathbf{X} - \tilde{\mathbf{X}}\|_F \leq \sqrt{\frac{1}{1 - \delta_{s+t}}} \|(\mathbf{X})_T\|_F + \frac{\sqrt{1 + \delta_t}}{1 - \delta_{s+t}} \|\mathbf{E}\|_F. \quad (18)$$

If  $t > s$ , define  $T_\nabla$  as the row-indices of the smallest  $t - s$  magnitude entries of  $\tilde{\mathbf{X}}$  in  $S$ , we have

$$\|\mathbf{X}_{T_\nabla}\|_F \leq \sqrt{2}\nu_2\delta_{s+t} \|\mathbf{X} - \tilde{\mathbf{X}}\|_F + \nu_2\sqrt{2(1 + \delta_t)} \|\mathbf{E}\|_F. \quad (19)$$

- First, we give an upper bound of  $\|\mathbf{X}_{T_\nabla}\|_F$ , by Lemma 6, let  $\mathbf{Z} = (\mathbf{W}_{T_0}\mathbf{X} - \mathbf{W}_{T_0}\tilde{\mathbf{X}})_S$ , we have

$$\begin{aligned} & \left\langle \mathbf{W}_{T_0}(\mathbf{X} - \tilde{\mathbf{X}}), \mathbf{A}^H \mathbf{A}(\mathbf{W}_{T_0}\mathbf{X} - \mathbf{W}_{T_0}\tilde{\mathbf{X}})_S \right\rangle \\ & + \left\langle \mathbf{W}_{T_0}\mathbf{E}, \mathbf{A}(\mathbf{W}_{T_0}\mathbf{X} - \mathbf{W}_{T_0}\tilde{\mathbf{X}})_S \right\rangle = 0. \end{aligned} \quad (20)$$

Noticing that  $\text{supp}(\tilde{\mathbf{X}}) \subseteq S$ , we have

$$\begin{aligned} & \left\| (\mathbf{W}_{T_0}\mathbf{X} - \mathbf{W}_{T_0}\tilde{\mathbf{X}})_S \right\|_F^2 \\ & = \left\langle \mathbf{W}_{T_0}(\mathbf{X} - \tilde{\mathbf{X}}), (\mathbf{W}_{T_0}\mathbf{X} - \mathbf{W}_{T_0}\tilde{\mathbf{X}})_S \right\rangle \\ & \stackrel{(20)}{=} \left\langle \mathbf{W}_{T_0}(\mathbf{X} - \tilde{\mathbf{X}}), \right. \\ & \quad \left. (\mathbf{I}_L - \mathbf{A}^H \mathbf{A})(\mathbf{X} - \tilde{\mathbf{X}})_S \right\rangle \\ & - \left\langle \mathbf{W}_{T_0}\mathbf{E}, \mathbf{A}(\mathbf{W}_{T_0}\mathbf{X} - \mathbf{W}_{T_0}\tilde{\mathbf{X}})_S \right\rangle \\ & \leq \stackrel{(7)}{\omega\delta_{s+t}} \left\| (\mathbf{W}_{T_0}\mathbf{X} - \mathbf{W}_{T_0}\tilde{\mathbf{X}})_S \right\|_F \left\| \mathbf{X} - \tilde{\mathbf{X}} \right\|_F \\ & \quad + \omega \|\mathbf{E}\|_F \sqrt{1 + \delta_t} \left\| \mathbf{W}_{T_0}(\mathbf{X} - \tilde{\mathbf{X}})_S \right\|_F. \end{aligned} \quad (21)$$

Divide both sides by  $\left\| (\mathbf{W}_{T_0}\mathbf{X} - \mathbf{W}_{T_0}\tilde{\mathbf{X}})_S \right\|_F$  to obtain (17).

- By expanding Lemma 2 in [25] to MMV model, we could get a relationship between  $\|\mathbf{X} - \tilde{\mathbf{X}}\|_F$  and  $\|\mathbf{X}_{\bar{S}}\|_F$ . We have

$$\|\mathbf{X} - \tilde{\mathbf{X}}\|_F \leq \sqrt{\frac{1}{1 - \delta_{s+t}}} \|\mathbf{X}_{\bar{S}}\|_F^2 + \frac{\sqrt{1 + \delta_t}}{1 - \delta_{s+t}} \|\mathbf{E}\|_F. \quad (22)$$

- Finally, we established the relationship between  $\mathbf{X}_{T_\nabla}$  and  $\mathbf{X} - \tilde{\mathbf{X}}$ . There exist a subset  $S' \subseteq S$  and  $S' \cap T = \emptyset$ . Since  $T_\nabla$  is defined by the set of indices of the  $t - s$  smallest row entries of  $\tilde{\mathbf{X}}$ , we can conclude that

$$\begin{aligned} & \left\| \tilde{\mathbf{X}}_{T_\nabla} \right\|_F + \left\| \mathbf{W}_{T_0}\tilde{\mathbf{X}}_{T_\nabla} \right\|_F \\ & \leq \left\| \tilde{\mathbf{X}}_{S'} \right\|_F + \left\| \mathbf{W}_{T_0}\tilde{\mathbf{X}}_{S'} \right\|_F. \end{aligned} \quad (23)$$

By eliminating the contribution from  $T_\nabla \cap S'$  and noticing that  $S' \cap T = \emptyset$ , we have

$$\begin{aligned} & \left\| \tilde{\mathbf{X}}_{T_\nabla \setminus S'} \right\|_F + \left\| \mathbf{W}_{T_0}\tilde{\mathbf{X}}_{T_\nabla \setminus S'} \right\|_F \\ & \leq \left\| (\tilde{\mathbf{X}} - \mathbf{X})_{S' \setminus T_\nabla} \right\|_F \\ & + \left\| \mathbf{W}_{T_0}(\tilde{\mathbf{X}} - \mathbf{X})_{S' \setminus T_\nabla} \right\|_F. \end{aligned} \quad (24)$$

For the left-hand side of (24), we have

$$\begin{aligned} & \left\| \tilde{\mathbf{X}}_{T_\nabla \setminus S'} \right\|_F + \left\| \mathbf{W}_{T_0}\tilde{\mathbf{X}}_{T_\nabla \setminus S'} \right\|_F \\ & = \left\| (\tilde{\mathbf{X}} - \mathbf{X} + \mathbf{X})_{T_\nabla \setminus S'} \right\|_F \\ & \quad + \left\| (\mathbf{W}_{T_0}(\tilde{\mathbf{X}} - \mathbf{X}) + \mathbf{W}_{T_0}\mathbf{X})_{T_\nabla \setminus S'} \right\|_F \\ & \geq \left\| \mathbf{X}_{T_\nabla} \right\|_F + \left\| \mathbf{W}_{T_0}\mathbf{X}_{T_\nabla} \right\|_F \end{aligned} \quad (25)$$

$$\begin{aligned} & - \left\| (\tilde{\mathbf{X}} - \mathbf{X})_{T_\nabla \setminus S'} \right\|_F \\ & - \left\| \mathbf{W}_{T_0}(\tilde{\mathbf{X}} - \mathbf{X})_{T_\nabla \setminus S'} \right\|_F. \end{aligned} \quad (26)$$

Combining (26) and (24), and noticing that

$$(T_\nabla \setminus S') \cap (S' \setminus T_\nabla) = \emptyset \quad (27)$$

$$(T_\nabla \setminus S') \cup (S' \setminus T_\nabla) \subseteq T, \quad (28)$$

we have

$$\begin{aligned} & \left\| \mathbf{X}_{T_\nabla} \right\|_F + \left\| \mathbf{W}_{T_0}\mathbf{X}_{T_\nabla} \right\|_F \\ & \leq \left\| (\tilde{\mathbf{X}} - \mathbf{X})_{T_\nabla \setminus S'} \right\|_F + \left\| (\mathbf{W}_{T_0}(\tilde{\mathbf{X}} - \mathbf{X}))_{T_\nabla \setminus S'} \right\|_F \\ & \quad + \left\| (\tilde{\mathbf{X}} - \mathbf{X})_{S' \setminus T_\nabla} \right\|_F + \left\| (\mathbf{W}_{T_0}(\tilde{\mathbf{X}} - \mathbf{X}))_{S' \setminus T_\nabla} \right\|_F \\ & \leq \sqrt{2} \left\| (\tilde{\mathbf{X}} - \mathbf{X})_{(S' \setminus T_\nabla) \cup (T_\nabla \setminus S')} \right\|_F \\ & \quad + \sqrt{2} \left\| (\mathbf{W}_{T_0}(\tilde{\mathbf{X}} - \mathbf{X}))_{(S' \setminus T_\nabla) \cup (T_\nabla \setminus S')} \right\|_F \\ & \leq \sqrt{2} \left\| (\tilde{\mathbf{X}} - \mathbf{X})_S \right\|_F + \sqrt{2} \left\| (\mathbf{W}_{T_0}(\tilde{\mathbf{X}} - \mathbf{X}))_S \right\|_F \\ & \stackrel{(17)}{\leq} \sqrt{2}(1 + \omega)\delta_{s+t} \left\| \mathbf{X} - \tilde{\mathbf{X}} \right\|_F \\ & \quad + (1 + \omega)\sqrt{2(1 + \delta_t)} \|\mathbf{E}\|_F. \end{aligned} \quad (29)$$

Also, we can obtain the relationship between  $\|\mathbf{W}_{T_0}\mathbf{X}_{T_\nabla}\|_F$  and  $\|\mathbf{X}_{T_\nabla}\|_F$ :

$$\eta\omega\sqrt{s_3}\|\mathbf{X}_{T_\nabla}\|_F \leq \|\mathbf{W}_{T_0}\mathbf{X}_{T_\nabla}\|_F. \quad (30)$$

Noting the definition of  $\nu_2$ , combining (29) and (30), we have

$$\begin{aligned} \|\mathbf{X}_{T_\nabla}\|_F &\leq \frac{\sqrt{2}(1+\omega)\delta_{s+t}}{1+\eta\omega\sqrt{s_3}} \|\mathbf{X} - \tilde{\mathbf{X}}\|_F \\ &\quad + \frac{(1+\omega)\sqrt{2(1+\delta_t)}}{1+\eta\omega\sqrt{s_3}} \|\mathbf{E}\|_F. \end{aligned} \quad (31)$$

We have completed the proof of Lemma 2.

**Remark 2.** When we consider the atom selection strategy of  $\|\tilde{\mathbf{X}}_{T_\nabla} + \mathbf{W}_{T_0}\tilde{\mathbf{X}}_{T_\nabla}\|_F \leq \|\tilde{\mathbf{X}}_{S'} + \mathbf{W}_{T_0}\tilde{\mathbf{X}}_{S'}\|_F$ , we can also obtain another upper bound for  $\|\mathbf{X}_{T_\nabla}\|_F$  in (19). In this case, we should allocate  $2\|\mathbf{X}_{T_\nabla}\|_F$  to the left hand side of (25), we have

$$\begin{aligned} \|\mathbf{X}_{T_\nabla}\|_F &\leq \sqrt{2}\nu_3\delta_{s+t} \|\mathbf{X} - \tilde{\mathbf{X}}\|_F \\ &\quad + \nu_4\sqrt{2(1+\delta_t)} \|\mathbf{E}\|_F. \end{aligned} \quad (32)$$

where  $\nu_3 = (1 - \omega + \omega\delta_{s+t} + \delta_{s+t})/(2\delta_{s+t})$  and  $\nu_4 = (1 + \omega)/(2\delta_{s+t})$ .

**Lemma 3.** In steps 4 and 5 of SI-SSP, we have

$$\|\mathbf{X}_{\bar{S}^k}\|_F \leq \sqrt{2}\varphi_1\delta_{3s} \|\mathbf{X} - \mathbf{X}^{k-1}\|_F + \varphi_1\sqrt{2(1+\delta_{3s})} \|\mathbf{E}\|_F. \quad (33)$$

*Proof:* From step 5 of SI-SSP, we have

$$\mathbf{X}_{S_i}^k = \arg \min_{\Theta: \text{supp}(\Theta) = S_{S_i}^k} \|\mathbf{Y}_{S_i} - \mathbf{A}\Theta\|_F. \quad (34)$$

From the step 4 of SI-SSP, let  $\mathbf{X}^k = [\mathbf{X}_{S_1}^k, \dots, \mathbf{X}_{S_r}^k]$ , we have the following conclusion

$$\begin{aligned} &\|(\mathbf{A}^H(\mathbf{Y} - \mathbf{A}\mathbf{X}^{k-1}))_T\|_F \\ &+ \|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{Y} - \mathbf{A}\mathbf{X}^{k-1}))_T\|_F \\ &\leq \|(\mathbf{A}^H(\mathbf{Y} - \mathbf{A}\mathbf{X}^{k-1}))_{\Delta S}\|_F \\ &+ \|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{Y} - \mathbf{A}\mathbf{X}^{k-1}))_{\Delta S}\|_F. \end{aligned} \quad (35)$$

By removing the same coordinates  $T \cap \Delta S$ , we can get

$$\begin{aligned} &\|(\mathbf{A}^H(\mathbf{Y} - \mathbf{A}\mathbf{X}^{k-1}))_{T \setminus \Delta S}\|_F \\ &+ \|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{Y} - \mathbf{A}\mathbf{X}^{k-1}))_{T \setminus \Delta S}\|_F \\ &\leq \|(\mathbf{A}^H(\mathbf{Y} - \mathbf{A}\mathbf{X}^{k-1}))_{\Delta S \setminus T}\|_F \\ &+ \|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{Y} - \mathbf{A}\mathbf{X}^{k-1}))_{\Delta S \setminus T}\|_F. \end{aligned} \quad (36)$$

Because  $\text{supp}(\mathbf{X}) = T$  and  $\text{supp}(\mathbf{X}^{k-1}) = S^{k-1}$ ,

$$(\mathbf{X} - \mathbf{X}^{k-1})_{\Delta S \setminus (T \cup S^{k-1})} = 0. \quad (37)$$

For the right-hand of (36), we have

$$\begin{aligned} &\|(\mathbf{A}^H(\mathbf{Y} - \mathbf{A}\mathbf{X}^{k-1}))_{\Delta S \setminus T}\|_F \\ &+ \|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{Y} - \mathbf{A}\mathbf{X}^{k-1}))_{\Delta S \setminus T}\|_F \end{aligned}$$

$$\begin{aligned} &= \|(\mathbf{A}^H(\mathbf{Y} - \mathbf{A}\mathbf{X}^{k-1}))_{\Delta S \setminus (T \cup S^{k-1})}\|_F \\ &+ \|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{Y} - \mathbf{A}\mathbf{X}^{k-1}))_{\Delta S \setminus (T \cup S^{k-1})}\|_F \\ &= \|(\mathbf{A}^H(\mathbf{A}\mathbf{X} + \mathbf{E} - \mathbf{A}\mathbf{X}^{k-1}))_{\Delta S \setminus (T \cup S^{k-1})}\|_F \\ &+ \|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{A}\mathbf{X} + \mathbf{E} - \mathbf{A}\mathbf{X}^{k-1}))_{\Delta S \setminus (T \cup S^{k-1})}\|_F \\ &\leq \|((\mathbf{A}^H\mathbf{A} - \mathbf{I}_L)(\mathbf{X} - \mathbf{X}^{k-1}))_{\Delta S \setminus T}\|_F \\ &\quad + \|(\mathbf{A}^H\mathbf{E})_{\Delta S \setminus T}\|_F \\ &+ \|(\mathbf{W}_{T_0}(\mathbf{A}^H\mathbf{A} - \mathbf{I}_L)(\mathbf{X} - \mathbf{X}^{k-1}))_{\Delta S \setminus T}\|_F \\ &\quad + \|(\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{E})_{\Delta S \setminus T}\|_F. \end{aligned} \quad (38)$$

Note that  $\tilde{S}^k = S^{k-1} \cup \Delta S$ , we have

$$(\mathbf{X} - \mathbf{X}^{k-1})_{T \setminus \tilde{S}^k} = \mathbf{X}_{\bar{S}^k}. \quad (39)$$

For the left-side of (36), we have

$$\begin{aligned} &\|(\mathbf{A}^H(\mathbf{Y} - \mathbf{A}\mathbf{X}^{k-1}))_{T \setminus \Delta S}\|_F \\ &+ \|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{Y} - \mathbf{A}\mathbf{X}^{k-1}))_{T \setminus \Delta S}\|_F \\ &= \|(\mathbf{A}^H(\mathbf{Y} - \mathbf{A}\mathbf{X}^{k-1}))_{T \setminus (\Delta S \cup S^{k-1})}\|_F \\ &+ \|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{Y} - \mathbf{A}\mathbf{X}^{k-1}))_{T \setminus (\Delta S \cup S^{k-1})}\|_F \\ &= \|(\mathbf{A}^H(\mathbf{A}\mathbf{X} + \mathbf{E} - \mathbf{A}\mathbf{X}^{k-1}))_{T \setminus (\Delta S \cup S^{k-1})}\|_F \\ &+ \|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{A}\mathbf{X} + \mathbf{E} - \mathbf{A}\mathbf{X}^{k-1}))_{T \setminus (\Delta S \cup S^{k-1})}\|_F \\ &= \|((\mathbf{A}^H\mathbf{A} - \mathbf{I}_L)(\mathbf{X} - \mathbf{X}^{k-1}) \\ &\quad + \mathbf{A}^H\mathbf{E} + \mathbf{X})_{\Delta S \setminus (T \cup S^{k-1})}\|_F \\ &\quad + \|\mathbf{W}_{T_0}(\mathbf{A}^H\mathbf{A} - \mathbf{I}_L)(\mathbf{X} - \mathbf{X}^{k-1}) \\ &\quad + \mathbf{W}_{T_0}\mathbf{A}^H\mathbf{E} + \mathbf{W}_{T_0}\mathbf{X})_{\Delta S \setminus (T \cup S^{k-1})}\|_F \\ &\geq \|\mathbf{X}_{\bar{S}^k}\|_F + \|(\mathbf{W}_{T_0}\mathbf{X})_{\bar{S}^k}\|_F \\ &\quad - \|((\mathbf{A}^H\mathbf{A} - \mathbf{I}_L)(\mathbf{X} - \mathbf{X}^{k-1}))_{\bar{S}^k}\|_F \\ &\quad - \|(\mathbf{W}_{T_0}(\mathbf{A}^H\mathbf{A} - \mathbf{I}_L)(\mathbf{X} - \mathbf{X}^{k-1}))_{\bar{S}^k}\|_F \\ &\quad - \|(\mathbf{A}^H\mathbf{E})_{\bar{S}^k}\|_F - \|(\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{E})_{\bar{S}^k}\|_F. \end{aligned} \quad (40)$$

Combining (42) and (41), we have

$$\begin{aligned} &\|\mathbf{X}_{\bar{S}^k}\|_F + \|(\mathbf{W}_{T_0}\mathbf{X})_{\bar{S}^k}\|_F \\ &\leq \|((\mathbf{A}^H\mathbf{A} - \mathbf{I}_L)(\mathbf{X} - \mathbf{X}^{k-1}))_{T \setminus \tilde{S}^k}\|_F \\ &+ \|(\mathbf{W}_{T_0}(\mathbf{A}^H\mathbf{A} - \mathbf{I}_L)(\mathbf{X} - \mathbf{X}^{k-1}))_{T \setminus \tilde{S}^k}\|_F \\ &+ \|(\mathbf{W}_{T_0}(\mathbf{A}^H\mathbf{A} - \mathbf{I}_L)(\mathbf{X} - \mathbf{X}^{k-1}))_{\Delta S \setminus T}\|_F \\ &\quad + \|((\mathbf{A}^H\mathbf{A} - \mathbf{I}_L)(\mathbf{X} - \mathbf{X}^{k-1}))_{\Delta S \setminus T}\|_F \\ &\quad + \|(\mathbf{A}^H\mathbf{E})_{T \setminus \tilde{S}^k}\|_F + \|(\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{E})_{T \setminus \tilde{S}^k}\|_F \\ &+ \|(\mathbf{A}^H\mathbf{E})_{\Delta S \setminus T}\|_F + \|(\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{E})_{\Delta S \setminus T}\|_F \\ &\leq \sqrt{2} \|((\mathbf{A}^H\mathbf{A} - \mathbf{I}_L)(\mathbf{X} - \mathbf{X}^{k-1}))_{T \cup \Delta S}\|_F \\ &\quad + \sqrt{2} \|(\mathbf{W}_{T_0}(\mathbf{A}^H\mathbf{A} - \mathbf{I}_L)(\mathbf{X} - \mathbf{X}^{k-1}))_{T \cup \Delta S}\|_F \\ &\quad + \sqrt{2} \|(\mathbf{A}^H\mathbf{E})_{T \cup \Delta S}\|_F + \sqrt{2} \|(\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{E})_{T \cup \Delta S}\|_F \\ &\leq \sqrt{2}(1+\omega)\delta_{3s} \|\mathbf{X} - \mathbf{X}^{k-1}\|_F \\ &\quad + (1+\omega)\sqrt{2(1+\delta_{3s})} \|\mathbf{E}\|_F. \end{aligned} \quad (42)$$

We can obtain the relationship between  $\|\mathbf{X}_{\bar{S}^k}\|_F$  and  $\|(\mathbf{W}_{T_0}\mathbf{X})_{\bar{S}^k}\|_F$ :

$$\eta\omega\sqrt{s_2}\|\mathbf{X}_{\bar{S}^k}\|_F \leq \|(\mathbf{W}_{T_0}\mathbf{X})_{\bar{S}^k}\|_F. \quad (43)$$

Noting the definition of  $\nu_1$ , combining (42) and (43), we have

$$\begin{aligned} \|\mathbf{X}_{\bar{S}^k}\|_F &\leq \frac{\sqrt{2}(1+\omega)\delta_{3s}}{1+\eta\omega\sqrt{s_2}}\|\mathbf{X}-\tilde{\mathbf{X}}\|_F \\ &\quad + \frac{(1+\omega)\sqrt{2(1+\delta_{2s})}}{1+\eta\omega\sqrt{s_2}}\|\mathbf{E}\|_F. \end{aligned} \quad (44)$$

We have completed the proof of Lemma 3.

**Remark 3.** When we consider the atom selection strategy in select step that

$$\begin{aligned} &\|((\mathbf{I}_L + \mathbf{W}_{T_0})\mathbf{A}^H(\mathbf{Y} - \mathbf{A}\mathbf{X}^{k-1}))_T\|_F \\ &\leq \|((\mathbf{I}_L + \mathbf{W}_{T_0})\mathbf{A}^H(\mathbf{Y} - \mathbf{A}\mathbf{X}^{k-1}))_{\Delta S}\|_F. \end{aligned} \quad (45)$$

We can also obtain another upper bound for  $\|\mathbf{X}_{\bar{S}^k}\|_F$  in (33). In this case, we should allocate  $2\|\mathbf{X}_{\bar{S}^k}\|_F$  to the left hand side of (40), we have

$$\begin{aligned} \|\mathbf{X}_{\bar{S}^k}\|_F &\leq \sqrt{2}\nu_4\delta_{3s}\|\mathbf{X}-\mathbf{X}^{k-1}\|_F \\ &\quad + \nu_4\sqrt{2(1+\delta_{3s})}\|\mathbf{E}\|_F. \end{aligned} \quad (46)$$

where  $\nu_4 = (1 - \omega + \omega\delta_{3s} + \delta_{3s})/(2\delta_{3s})$  and  $\nu_4 = (1 + \omega)/(2\delta_{3s})$ . Based on conclusions (32) and (46), we know that the sensing matrix  $\mathbf{A}$  obeys the RIP with

$$\delta_{3s} \leq \sqrt{\frac{\nu_3\sqrt{\nu_3^2 + 4\nu_4^2} - \nu_3^2 - 1}{4\nu_3^2\nu_4^2 - 2\nu_3^2 - 1}}. \quad (47)$$

**Lemma 4.** Let  $T_0 \subseteq \{1, 2, \dots, n\}$ , for two vectors  $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ , if  $|\text{supp}(\mathbf{u}) \cup \text{supp}(\mathbf{v})| \leq t$ ,

$$|\langle \mathbf{u}, (\mathbf{W}_{T_0} - \mathbf{W}_{T_0}\mathbf{A}^H\mathbf{A})\mathbf{v} \rangle| \leq \omega\delta_t\|\mathbf{u}\|\|\mathbf{v}\|; \quad (48)$$

moreover, if  $U \subseteq \{1, 2, \dots, n\}$  and  $|U \cup \text{supp}(\mathbf{v})| \leq t$ , then

$$|(\mathbf{W}_{T_0} - \mathbf{W}_{T_0}\mathbf{A}^H\mathbf{A})\mathbf{v}| \leq \omega\delta_t\|\mathbf{v}\|. \quad (49)$$

*Proof:* the RIC  $\delta_s$  can be expressed as [25]

$$\delta_s = \max_{S \subseteq \{1, 2, \dots, N\}, |S| \leq s} \|\mathbf{A}_S^*\mathbf{A}_S - \mathbf{I}\|_{2 \rightarrow 2}, \quad (50)$$

where

$$\|\mathbf{A}_S^*\mathbf{A}_S - \mathbf{I}\|_{2 \rightarrow 2} = \sup_{\mathbf{a} \in \mathbb{R}^{|S|} \setminus \{0\}} \frac{\|(\mathbf{A}_S^*\mathbf{A}_S - \mathbf{I})\mathbf{a}\|_2}{\|\mathbf{a}\|_2}. \quad (51)$$

Let  $S = \text{supp}(\mathbf{u}) \cup \text{supp}(\mathbf{v})$ , then  $|S| \leq t$ . Let  $\mathbf{u}_{|S}, \mathbf{v}_{|S}$  denote respectively the  $S$ -dimensional sub-vectors of  $\mathbf{u}$  and

$\mathbf{v}$  obtained by only keeping the components indexed by  $S$ . We have

$$\begin{aligned} &|\langle \mathbf{u}, (\mathbf{W}_{T_0} - \mathbf{W}_{T_0}\mathbf{A}^H\mathbf{A})\mathbf{v} \rangle| \\ &= |\langle \mathbf{W}_{T_0}\mathbf{u}, \mathbf{v} \rangle - \langle \mathbf{A}\mathbf{W}_{T_0}\mathbf{u}, \mathbf{A}\mathbf{v} \rangle| \\ &= |\langle \mathbf{W}_{T_0}\mathbf{u}_{|T}, (\mathbf{I}_L - \mathbf{A}_T^H\mathbf{A}_T)\mathbf{v}_{|T} \rangle| \\ &\leq \|\mathbf{W}_{T_0}\mathbf{u}_{|T}\|_2 \|(\mathbf{I}_L - \mathbf{A}_T^H\mathbf{A}_T)\mathbf{v}_{|T}\|_2 \\ &\stackrel{(51)}{\leq} \|\mathbf{W}_{T_0}\mathbf{u}_{|T}\|_2 \|\mathbf{I}_L - \mathbf{A}_T^H\mathbf{A}_T\|_{2 \rightarrow 2} \|\mathbf{v}_{|T}\|_2 \\ &\stackrel{(50)}{\leq} \omega\delta_t\|\mathbf{u}_{|T}\|_2 \|\mathbf{v}_{|T}\|_2 \\ &= \omega\delta_t\|\mathbf{u}\|_2\|\mathbf{v}\|_2, \end{aligned} \quad (52)$$

moreover, we have

$$\begin{aligned} &\|((\mathbf{W}_{T_0} - \mathbf{W}_{T_0}\mathbf{A}^H\mathbf{A})\mathbf{v})_U\|_2^2 \\ &= \langle ((\mathbf{W}_{T_0} - \mathbf{W}_{T_0}\mathbf{A}^H\mathbf{A})\mathbf{v})_U, (\mathbf{W}_{T_0} - \mathbf{W}_{T_0}\mathbf{A}^H\mathbf{A})\mathbf{v} \rangle \\ &\stackrel{(48)}{\leq} \delta_t \|((\mathbf{W}_{T_0} - \mathbf{A}^H\mathbf{A})\mathbf{v})_U\|_2 \|\mathbf{v}\|_2. \end{aligned} \quad (53)$$

which completes the proof of Lemma 4.

**Lemma 5.** For SMV model  $\mathbf{y} = \Phi\mathbf{x} + \mathbf{e}$ , let  $T_0 \subseteq \{1, 2, \dots, n\}$ , let  $U \subseteq \{1, 2, \dots, n\}$  and  $|U \cap T_0| \leq u$ , we have

$$\|(\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{e})_U\|_2 \leq \omega\delta_u\|\mathbf{e}\|_2. \quad (54)$$

*Proof:* The lemma easily follows from the fact that

$$\begin{aligned} &\|(\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{e})_U\|_2^2 \\ &= \langle \mathbf{W}_{T_0}\mathbf{A}^H\mathbf{e}, (\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{e})_U \rangle \\ &= \langle \mathbf{e}, \mathbf{W}_{T_0}\mathbf{A}((\mathbf{A}^H\mathbf{e})_U) \rangle \\ &\leq \|\mathbf{e}\|_2 \|\mathbf{W}_{T_0}\mathbf{A}((\mathbf{A}^H\mathbf{e})_U)\|_2 \\ &\stackrel{(3)}{\leq} \|\mathbf{e}'\|_2 \omega\sqrt{1+\delta_u} \|(\mathbf{A}^H\mathbf{e})_U\|_2. \end{aligned} \quad (55)$$

**Lemma 6.** Consider MMV model  $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{E}$ , let  $\tilde{\mathbf{X}}$  be the solution of the least squares problem  $\arg \min_{\mathbf{Z}} \{\|\mathbf{Y} - \mathbf{A}\mathbf{Z}\|_F, \text{supp}(\mathbf{Z}) \subseteq S\}$ , then

$$\langle \mathbf{W}_{T_0}\mathbf{X} - \mathbf{W}_{T_0}\tilde{\mathbf{X}}, \mathbf{A}^H\mathbf{A}\mathbf{Z} \rangle + \omega \langle \mathbf{E}, \mathbf{A}\mathbf{Z} \rangle = 0 \quad (56)$$

*Proof:* Due to the orthogonality, the residue  $\mathbf{Y} - \mathbf{A}\tilde{\mathbf{X}}$  is orthogonal to the space  $\mathbf{A}\mathbf{Z}, \text{supp}(\mathbf{Z}) \subseteq S$ . This means that for all  $\mathbf{Z} \in \mathbb{R}$  with  $\text{supp}(\mathbf{Z}) \subseteq S$ ,

$$\langle \mathbf{A}(\mathbf{Y} - \mathbf{A}\tilde{\mathbf{X}}), \mathbf{Z} \rangle = 0 \quad (57)$$

then, let  $\tilde{\mathbf{X}}'$  be the solution of the least squares problem  $\arg \min_{\mathbf{Z}} \{\|\mathbf{Y}' - \mathbf{A}\mathbf{Z}\|_F, \text{supp}(\mathbf{Z}) \subseteq S\}$ , where  $\mathbf{Y}' = \frac{\mathbf{A}\mathbf{W}_{T_0}\mathbf{X}_{T_0}}{\omega} + \mathbf{E}$ . We have

$$\tilde{\mathbf{X}}' = \frac{\mathbf{W}_{T_0}\tilde{\mathbf{X}}}{\omega}. \quad (58)$$

Then, by (57), we have

$$\begin{aligned} 0 &= \left\langle \frac{\mathbf{A}\mathbf{W}_{T_0}\mathbf{X}_{T_0}}{\omega} + \mathbf{E} - \mathbf{A}\frac{\mathbf{W}_{T_0}\tilde{\mathbf{X}}}{\omega}, \mathbf{A}\mathbf{Z} \right\rangle \\ &= \langle \mathbf{W}_{T_0}\mathbf{X} - \mathbf{W}_{T_0}\tilde{\mathbf{X}}, \mathbf{A}^H\mathbf{A}\mathbf{Z} \rangle + \omega \langle \mathbf{E}, \mathbf{A}\mathbf{Z} \rangle. \end{aligned} \quad (59)$$

### B. Proof of Theorem 1

Firstly, we assume that a multi-coset sampler has  $p$  channels, and the sampling rate of each channel is determined by the sampled signal sequence, which is in the form of

$$x_{c_i}[n] = x(LTn + \tau_i), \quad n = 0, 1, \dots \quad (60)$$

The average sampling rate of the  $p$ -channel is  $p$  times that of one channel. Note that when the sampling rate of the  $p$ -channel is greater than the Nyquist rate, we only need to operate the Nyquist frequency to sample, the actual sampling rate is in the form of

$$\min\left(\frac{p}{LT}, f_{\text{nyq}}\right). \quad (61)$$

The theoretical lower bound of the sampling rate is given in [17], which is directly determined by the true bandwidth of the signal:

$$\min(2\lambda(\mathcal{T}), f_{\text{nyq}}). \quad (62)$$

In most cases, the subband bandwidth  $\lambda(\mathcal{T})$  and the actual sampling rate does not exceed  $f_{\text{nyq}}$  (when not satisfied, the sampling rate is  $f_{\text{nyq}}$ ). To ensure reconstruction performance, the parameter  $p$  is often not set too low (for instance, it's often chosen to be at least twice the sparsity level  $\text{supp}(\mathbf{X})$ ). It will be seen from above that the theoretical lower bound on the sampling rate is achieved only when  $p/LT = pf_s \leq 2\lambda(\mathcal{T}) = 2N_{\text{sig}}B$ . In other words, when  $p = 2\text{supp}(\mathbf{X})$  for the worst case of  $p$ , the condition for the actual sampling rate to meet the theoretical lower bound is  $K \leq \frac{N_{\text{sig}}B}{f_s}$ .

### C. Proof of Theorem 2

Consider all blocks  $\{\mathbf{X}^{U_1}, \dots, \mathbf{X}^{U_M}\}$  in  $\mathbf{X}$ , where there are consecutive corresponding frequency points of length  $B$ . For the case  $B > f_s$  and each block occupies  $M$  rows in  $\mathbf{X}$ , the length of the frequency point in each  $\bar{\mathbf{X}}^{U_i}$  meets

$$l = B - (M - 2)f_s < f_s. \quad (63)$$

Because  $l < f_s$ , we know that the non-zero elements of an any sub-block (i.e.  $\bar{\mathbf{X}}^{U_i}$ ) are distributed on both sides and do not intersect on the columns. Consider one block  $\bar{\mathbf{X}}^{U_i}$ . We let  $r \rightarrow \infty$  and observe the change in  $\bar{\mathbf{X}}^{U_i}$ .  $\bar{\mathbf{X}}^{U_i}$  gradually changes from an MMV form to an SMV form. In SMV, the sparsity of the signal is determined by the column in  $\bar{\mathbf{X}}^{U_i}$  where it is located. It is observed that

$$\lim_{r \rightarrow \infty, i, j} \text{supp}(\bar{\mathbf{X}}_{:,j}^{U_i}) \leq 1. \quad (64)$$

Thus, the sparsity of each column in  $\bar{\mathbf{X}}^{U_i}$  is less than  $N_{\text{sig}}$ .

A more complex situation is when  $r$  is a finite value, assuming  $r = r^*$  is a finite value. In this case, the length of frequency point in each sub-matrix is  $\frac{f_s}{r^*}$ . We use reduction to absurdity to prove the condition that the sparsity of each sub-matrix is less than  $N_{\text{sig}}$ . Assuming that there exists a sub-matrix  $\bar{\mathbf{X}}_{S_i}$  with  $\text{supp}(\bar{\mathbf{X}}_{S_i}) > N_{\text{sig}}$ . Also, the non-zero elements of an any sub-block of  $\bar{\mathbf{X}}_{S_i}$  do not intersect on the columns.  $\bar{\mathbf{X}}_{S_i}$  must contain both non-zero elements on both

sides of one block. From (63), we know that the length of any block of  $\bar{\mathbf{X}}$  is less than  $f_s$ . We can draw a conclusion that

$$\frac{f_s}{r^*} > f_s - l = (M - 1)f_s - B. \quad (65)$$

As can be seen, there exist a contradiction between (65) and Theorem 2, so the length of any column-partition sub-matrix non-zero elements in  $\bar{\mathbf{X}}$  must be less or equal than  $f_s - l$ , which is equivalent to  $r$  satisfying

$$r \geq \lceil \frac{f_s}{(M - 1)f_s - B} \rceil. \quad (66)$$